Web100 Tools

Matt Mathis
John Heffner
Raghu Reddy
1/2/2003

info@web100.org

Submitted to CCR
Overview

- This presentation is based on a paper currently under review
  - Look for it in Computer Communications Review

- The problem

- Web100 Components

- Web100 to Facilitate Diagnostics

- Web100 to Facilitate Research
Why? TCP "tuning" requires experts

- TCP/IP hides the net from upper layers

- This is good for the growth of the ’net
  - but it hides all problems

- All bugs have the same symptom: less than expected performance!
TCP tuning is really debugging

- Six classes of bugs limit performance
  - TCP sender or receiver buffer space
  - Packet losses, corruption, congestion, lame HW
  - IP Routing, long round trip times
  - Inefficient applications
  - IP Packet reordering
  - Improper MSS negotiations or MTU discovery
TCP tuning is painful debugging

- Any one problem can mask all other problems confounding all but the best experts

- Akin to finding the weakest link of an invisible chain

- We need better diagnostic visibility
When there is a problem, just ask TCP
- TCP has an ideal vantage point
- TCP can identify the bottleneck subsystem
- TCP already measures the network
- TCP can measure the application
- TCP can adjust itself

See: www.web100.org
Web100 Components

- Kernel Instrument Set (KIS)
- Diagnostic Tools
- Autotuning
- Widely distributed Open Source
- TCP ESTATS MIB
- Promote vendor adoption
Kernel Instrument Set (KIS)

- TCP Instruments prototyped in Linux 2.4
- Simple API via /proc

Instrument groups
- Options and State
- IP Traffic and Throughput
- Triage
- Congestion Events
- Network Path properties
- TCP API Usage
- TCP Parameters
- WAD Parameters
Diagnostic Tools

- User mode tools to prove core functionality
- Template for future tool developers
- KIS validation
- Portable library to hide OS details
TCP Autotuning

- Superceeds today's controls for experts
  - Eliminate primary cause of the wizard gap
- New TCP buffer management model
- TCP just gets it right without controls
- Paper in draft for future publication
Software Distribution

- Open source
  - Linux kernel patches
  - User mode Tools
  - Contributed Software

- Active user/developer support by the web100 team
  - Really good feedback

- User group meetings
  - Really good feedback
IETF Standards

- TCP extended statistics MIB
  - Adds detailed per connection statistics

- Standard TCP instruments and controls

- Next big draft due March 3rd
  - TSV WG work item

- Key vendors are already participating
Promote vendor adoption

- Ultimate goal: TCP instruments in every OS as shipped
- Direct promotion to vendors
- Future Linux mainline inclusion
- Future RFQ language
Web100 Diagnostics

- Web100 basic tool suite
- Diagnostic servers
- Active path diagnosis
- Network data mining
Basic tools

- GTK widget based
  - Modular, and the parts are designed to be reused

- Assorted traffic instruments

- Triage
  - Reports why and how long TCP stopped sending
The Triage tool

- In this example the receiving application stalls
Diagnostic servers

- A server that performs simple diagnostics
  - Client is simple a browser, telnet, etc

- Java based clients measure the path
  - Rich Carlson @ ANL
  - Tom Dunigan @ ORNL

- Raghu Reddy’s SYN tester
  - http://syntest.psc.edu:7961/
  - telnet syntest.psc.edu 7960

- Huge promise for more developments
  - Only scratched the surface
Telnet syntest.psc.edu 7960

Your IP address is: 131.94.191.210

Check of SYN options

Variable : Val : Warning (if any)

SACKEnabled : 3 :
TimestampsEnabled : 1 :
CurMSS : 1448 :
WinScaleRcvd : 5 :
CurRwinRcvd : 182 :

End of SYN options

Please visit
http://www.psc.edu/networking/perf_tune.html
Active path diagnosis

- Test to see if a sub-path can support a longer path
  - Can my office to the GigaPoP support my office to Timbuktu?

- Still under development
  - [http://www.psc.edu/~web100/pathprobe/](http://www.psc.edu/~web100/pathprobe/)

- Future paper TBD
Network data mining

- Active: collect stats from periodic tests
  - Net100 NTAF
    - http://dpsslx05.lbl.gov/WK/NTAF/

- Passive: collect stats from normal traffic
  - Data mine/post process for diagnoses
Web100 Experimentation and Research

- Work-arounds
- Floyd’s High Speed TCP
- User Mode congestion control
Work-arounds and manual controls

- Techniques to coddle buggy gear

- Safe examples
  - Clamp cwnd to prevent full utilization
  - Clamp Slow Start to prevent queue overflow [Floyd]

- Unsafe examples
  - Defeat congestion control

- All partially replace standard automatic operation with context specific "manual" controls
User Mode congestion control

- Implements Vegas, etc in user mode

- Use WAD to replace default +1 additive increase
  - Disable default additive increase
  - Writing WAD_CwndAdjust causes cwnd adjustment
  - No change to multiplicative decrease

- User mode tool monitors RTT
  - Adjusts cwnd +/- 1 per alternate RTT

- Partition the problem
  - 8 lines of easy kernel code
  - All the complexity in easy to debug user mode
Anti-Conclusion

- The Web100 project has provided a core technology and prototypes for the next tier of user friendly network diagnostics.

- There is a vast fronter of unexplored diagnostic approaches

- I believe that Web100 can "easily" close the wizard gap by 2 orders of magnitude
  - Trivial to get 100 Mb/s from any university netport
  - Easy to get several hundred Mb/s